The Development of "Roadway Name" Table for the Idaho Transport Department's Milepost And Coded Segment (MACS) System

Final Report
February 2003

ITD contract – 0010 (027) RP (159)

Prepared for
Idaho Department of Transportation

Prepared by

BOISE STATE UNIVERSITY

Mandar Khanal

The Development of "Roadway Name" Table for the Idaho Transport Department's Milepost And Coded Segment (MACS) System

Final Report
February 2003

ITD contract – 0010 (027) RP (159)

Prepared for
Idaho Department of Transportation

Prepared by

BOISE STATE UNIVERSITY

Mandar Khanal

TABLE OF CONTENTS

1. INTRODUCTION 1

1.1 Background 1

1.2 Scope of Work 2

1.3 Problem Definition 3

1.4 Project Tasks 3

1.5 Project Deliverables 5

1.6 Organization of Report 6

2.0 THE CONFLATION TOOL 7

2.1 Introduction 7

2.2 General steps 8

2.3 Functions of The Conflation Tool 14

2.3.1 Buffer zone 14

2.3.2 Match 14

2.3.3 Make 16

2.3.3.1 B01 Control Table 16

2.3.3.2 B05 County Limits Table 17

2.3.3.3 B11 Jurisdiction Table 17

2.3.3.4 B16 RoadwayNames_highways table 19

2.3.4 C01Make 21

3. PERFORMANCE 25

4. TOOL VALIDATION 30

4.1 Sample Selection 30

4.1.1 Sample Size 30

4.2 Validation of Methodology 32

4.3. Field Observations 32

1st Street North Extended 32

4th and 5th Streets 33

10th Ave S 34

W. Blaine Ave 34

Chicago St, Ventura Dr, Wildflower Dr and Sunflower Dr 34

Greenhurst Rd 34

E Roosevelt Rd 35

Sugar St 35

4.4 Validation Summary 35

5. CONCLUSIONS 36

Future Work 36

Appendix A 37

Appendix B 38

REFERENCES 44

LIST OF FIGURES

Figure 2.1 The Conflation Tool 7

Figure 2.2 Sandpoint Road Map Before Matching 9

Figure 2.3 Buffer Zone of 10 meters Around ITD Segments (Solid Lines) 10

Figure 2.4 Zoomed-In View of Buffer Zone of 10 Meters 11

Figure 2.5 Map After 10-Meter Match Iteration 12

Figure 2.6 Map After 150-Meter Match Iteration 12

Figure 2.7 B01 Control Table 16

Figure 2.8 B05 County Limits Table 17

Figure 2.9 B11 Jurisdiction Table 18

Figure 2.10 B16Roadway Names Table 21

Figure 2.11 Spatial Intersection Tool Dialogue Box 22

Figure 2.12 C1MP Descriptions Table 24

LIST OF TABLES

Table 3.1 The Number Of Mismatches.. 26

Table 3.2 Performance Based On The Percentage Of Tiger Records Matched................ 28

Table 3.3 Distribution Of Mismatches... 29

Table 4.1 List Of Validation Sample 31

1. INTRODUCTION

1.1 Background

The Idaho Transportation Department (ITD) requires the development of a Federal Geographic Data Committee (FGDC) compliant base map for use with all Geographic Information System (GIS)-enabled applications throughout the department. Such applications include:

- Safety Analysis
- Corridor Analysis
- Travel Demand Analysis
- Traffic Forecasting

The FGDC has issued guidelines for the minimum data requirements necessary for an adequate GIS base map. These minimum data requirements are referred to collectively as the "framework" layers because they represent the minimum data requirements necessary to provide an adequate framework for any kind of spatial data collection, management, analysis and reporting (including pavement, congestion, bridges, signing, railroad crossings, maintenance, project programming, traffic and safety). Transportation is one of the "framework" layers.

The framework transportation data layer includes the centerlines of roads, trails and navigable waterways, ports and airports; bridges and tunnels; and will have the following attributes:

- Feature identification code using a linear referencing system
- Names (including route numbers)
- Street address range (where applicable)

- Type (Street, Road, Circle, Parkway, Blvd, Way etc.)
- Functional class (where applicable)

The linear location referencing system used at ITD is called the Milepost And Coded Segment/ Roadway Segment Database System (MACS) (*1*). The intent of the MACS system is to identify and reference the location of all roads in Idaho open for use by the public, hence the necessity of this project. The database key used by the system consists of segment codes, reference points, and dates (*2*). Each record in the database corresponds directly with a road or road segment that physically exists in reality.

The various tables in the MACS system contain attribute data that can be cross-referenced with each other by the database key. The Roadway Names Table is one such table in MACS that will facilitate the integration of other planning, maintenance, traffic, construction and safety data with roadway characteristics, functional class, federal aid funding, local road inventory, pavement management, sign inventory, site analysis, local access, cost accounting, and relative geographic location data.

The Roadway Names Table in MACS is designed to include names (including route numbers), address ranges, and street type. Thus, a base map with features attributed according to the MACS system would adequately meet the minimum FGDC standards.

The Roadway Names in MACS is, however, currently unpopulated. To fulfill this need ITD awarded a project to Boise State University (BSU) to research and develop the Roadway Names Table and other related tables needed by MACS. This report describes how BSU accomplished this task. Specifically, it explains how spatial features from ITD maps were conflated with name and address range attributes from an external source to create an FGDC minimum standards compliant base maps for Idaho's urban areas.

1.2 Scope of Work

There were 25 urban areas of population 5000 or more in 2001, the year when this project started. ITD had roadway maps for these areas. These maps were used for the spatial data. For attribute data related to road names and address ranges the Topologically Integrated Geographic Encoded Reference (TIGER®) data from the US Census Bureau was used (*3*).

1.3 Problem Definition

The spatial data was obtained from ITD in the form of their urban area maps. ITD does not have address range information for the spatial features representing the road network. ITD did have road names but the naming did not conform to the National Emergency Number Association (NENA) data standards. (*4*)

It is possible that the missing attribute information could be obtained from a private source, but it was not clear that ITD would be able to provide the maps and the database to the public free of cost. Getting the information from a government source, such as the US Census Bureau, on the other hand, would not preclude such use. Hence TIGER ® data was used.

With data coming from two different sources, the problem was to devise a way to merge the two. ITD uses the software, GeoMedia ™ (*5*), for its GIS work. GeoMedia ™ does not have the capability to conflate attribute information from TIGER ® with spatial features data from ITD. It is believed that no other GIS software in the market has this capability. Hence a conflation tool to fulfill this need was developed.

1.4 Project Tasks

The Cooperative Transportation Research Program Agreement between ITD and BSU identified the following tasks for this project. The agency responsible for the task is also listed.

#1. ITD- Research and document availability and appropriateness [and costs] of digital data sources. The single most comprehensive data set of roadway name and address range information is the Topologically Integrated Geographic Encoding and Referencing (TIGER ®) data produced by the US Census Bureau. ITD's GIS Section has acquired and will provide the latest available edition (TIGER ® 2000) of this data set. If necessary, and where available and cost effective, TIGER ® data will be supplemented with data sets acquired by ITD from road jurisdictions and other public and private sources.

#2. BSU- Using appropriate GIS desktop software, combine ITD Urban Area Segment Code maps with TIGER ® data. Research data conversion algorithm for

- Conflating ITD and TIGER ® map layers (graphics line work), including
- Matching ITD' s LRS (Milepost And Coded Segment (MACS) system) with TIGER ® street names and address ranges
- Assigning, where needed, MACS LRS PK values to named streets in TIGER ®
    - Use 100.000 as Beginning Mile Point (BegMP)
    - Create B01Conrol, B05 County Limits, B11 Jurisdiction & C01 MP Descriptions records that correspond to B16Roadyway Names. ITD will provide Table Structure and coding scheme for B01, B05, B11 and C01.
- Retaining conflated map layers.

#3. BSU- Populate the MACS "Street Names" table with the following attributes

- LRS PK values
- Urban/City Federal Information Processing Standards (FIPS) Location codes
- Street Names (e.g. Broadway, Van Buren)
- Address ranges for Right and Left sides of each named street (e.g., BegAddRt= 100, EndAddRt=190 & BegAddLt= 105, EndAddLt= 195)

- Prefixes (e.g., North (N), Southwest (SW), etc.)
- Suffix1 (e.g. North (N), Southwest (SW), etc.)
- Suffix2 (e.g., Ave., Rd., Way, Blvd, Str., etc.)
- Known Aliases (e.g., Old US95; maximum of three Aliases)
- Record Sequence Numbers (i.e., where more than one MACS record is required to represent an entire, homogeneously named street)

#4. BSU- Document all processes and work flows to complete project.

#5. ITD- QA/QC: research and verify samples of TIGER ® Street Names and address ranges data using published street name Atlas's and Index' s (e.g., dtG MAPS 2000 Edition of *Detail Map Pages and Street Index* for ADA County), and field checks.

1.5 Project Deliverables

- The following MACS Tables in MS ACCESS format
    - Roadway Name Table populated as outlined above (B16)
    - Control Table (B01)
    - County Limits Table (B05)
    - Jurisdiction Table (B11)
    - Mile Point Descriptions Table (C01)
- Line work map layers conflated from TIGER ® and ITD graphics data
- Data Conversion Algorithms for periodic updating and maintenance
- Maintenance Work Flows for periodic updating and maintenance

- Project documentation

- Final report in both camera-ready hard copy and on electronic media. (WORD 6.0 or later)

1.6 Organization of Report

The report begins with this introductory section. Section 2 will explain in detail the conflation tool that was developed to complete the required project tasks. The use of the tool is also described in depth.

Section 3 will present some performance measures for the tool. Field verification results are provided in Section 4. The main report ends with conclusions in Section 5.

There are two Appendices to this report. Appendix A provides instructions on installing the Conflation tool into GeoMedia. Appendix B provides step-by-step instructions on the use of the tool.

2.0 THE CONFLATION TOOL

2.1 Introduction

The Conflation tool was developed as a part of this project to complete various tasks listed above. It is a tool that can execute many operations. In the following sections those operations and the way this tool executes them are described. The dialog box for this tool is shown in Figure 2.1.

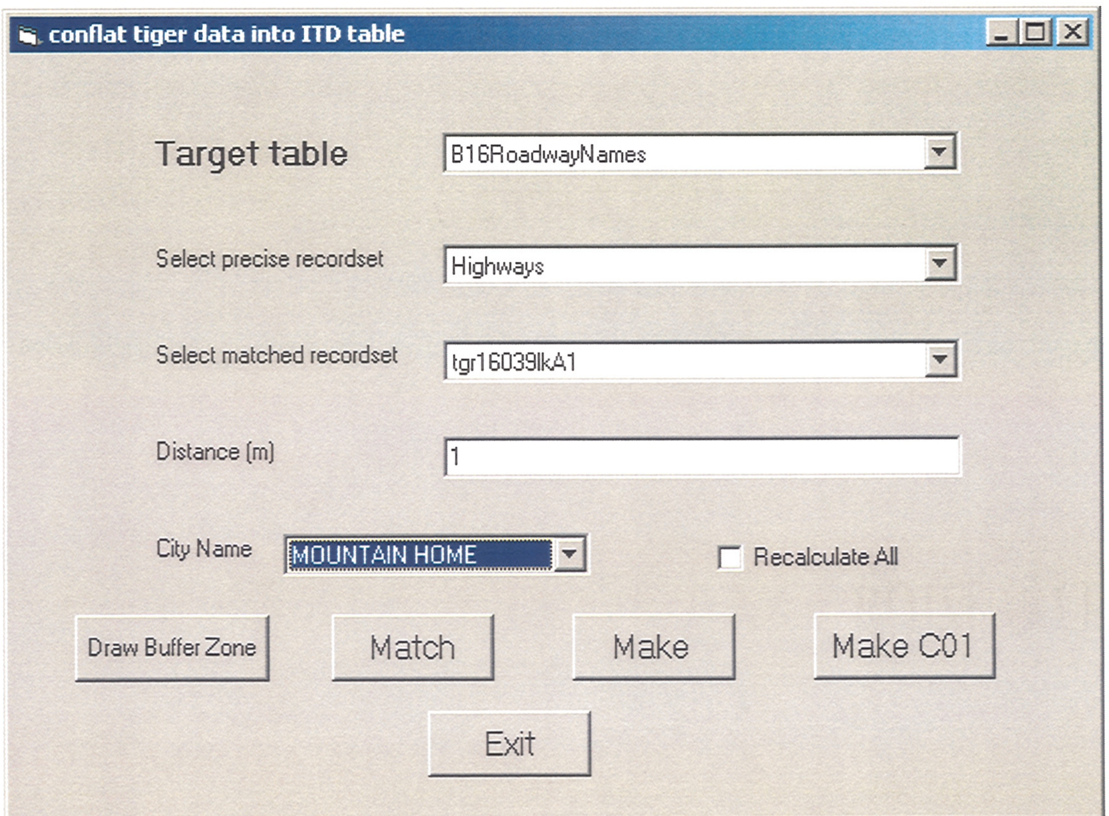


Figure 2.1 The Conflation Tool

Hypothesis

The following are the assumptions based on which the tool was designed.

1. The TIGER ® attribute data is correct but the map of linear features is not.

2. The record in the TIGER ® table that is the closest one to a record in the ITD table is matched to that record in the ITD table.

3. All records in TIGER ® that match to one record in the ITD table must have the same TIGER ® road name.

Implementation issues related to the second hypothesis:

If a TIGER segment has no name, it never matches to any record in ITD table. (But we can manually add it if it indeed matches to one record in ITD table.)

We deal with the FEDIRP and FEDIRS conflict check very strictly.

If all of the TIGER ® records that match to the same record in ITD table have the same FEDIRP, we say conflict_P is False. Otherwise, conflict_P is true.

If all of the records that match to the same record in ITD table have the same FEDIRS, we say conflict_S is False. Otherwise, conflict_S is true

We can therefore check for mistakes in the conflict_P, conflict_S fields and the map window.

2.2 General steps

1. Match as many records as possible from TIGER ® to the ITD table. Some manual work may be required in this step to deal with mismatch problems.

2. Using a new copy of the TIGER ® database and the ITD table that contains the match information, the B16 table is created.

3. The SeqNo field in the B16 table is filled manually to deal with roads that are represented by multiple segments.

In the following pages an example of the "matching" operation for the city of Sandpoint is described. Various features related with the process are also described with the help of screen captures from the GeoMedia software.

Roadway line features for the city of Sandpoint are shown in Figure 2.2. Solid lines in the map represent ITD segments while dashed lines represent TIGER ® segments

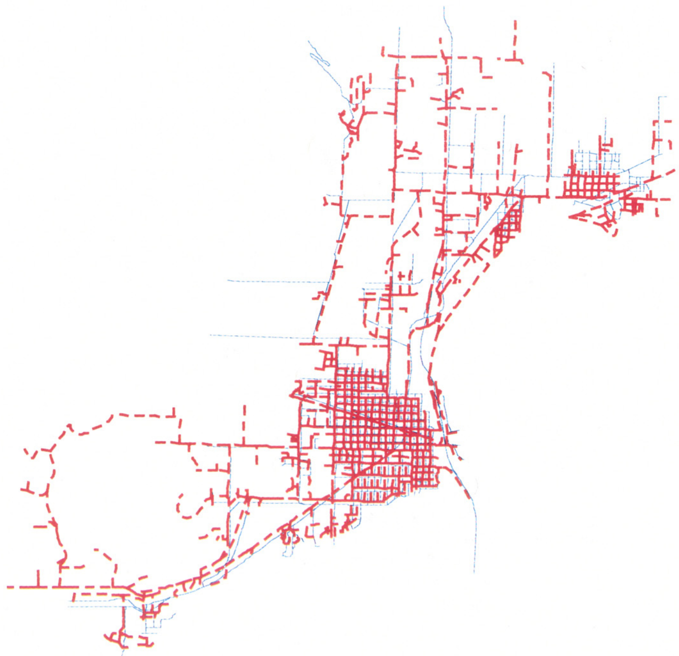


Figure 2.2 Sandpoint Road Map Before Matching

Two sets of features are depicted corresponding to the two sources, ITD and TIGER ®. The two sets are not aligned exactly. The ITD road segments are also longer than that of TIGER ®. There are only 196 ITD features corresponding to 1283 features from TIGER ® for this city. For any given ITD road segment there will be more than one corresponding TIGER ® features.

As can be seen from the figure, there are some line features from one source that do not have a corresponding set of features from the other source. In most cases there were many TIGER ® features that did not have a corresponding set of ITD features. But there were also a few cases in which TIGER ® did not have any features that could possibly match ITD features.

The first task in the development of the tool was to match the two sets of features. In other words, the number of TIGER ® features matching to one ITD feature had to be determined. This was accomplished by drawing a buffer zone around each ITD feature and including those TIGER ® features in the matching set that were contained totally within the buffer zones. Figure 2.3 shows 10-meter buffer zones drawn by the tool around every ITD feature.

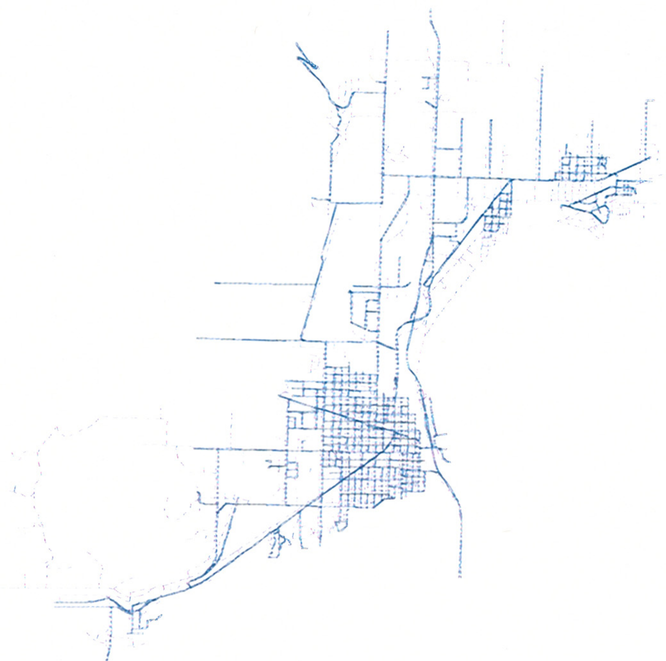


Figure 2.3 Buffer Zone of 10 meters Around ITD Segments (Solid Lines)

Figure 2.4 shows a larger scale map of a portion of the map shown in Figure 2.3. Dashed lines are the buffer zone around the solid ITD lines; TIGER ® segments are not shown in this figure.

The tool allows the user to select the width of the buffer zone, starting from one meter to any desired width. It is important to start the matching process with a narrow buffer zone. A wider buffer zone can contain many TIGER ® records and thus will complicate the task of finding the ones that match the ITD record.

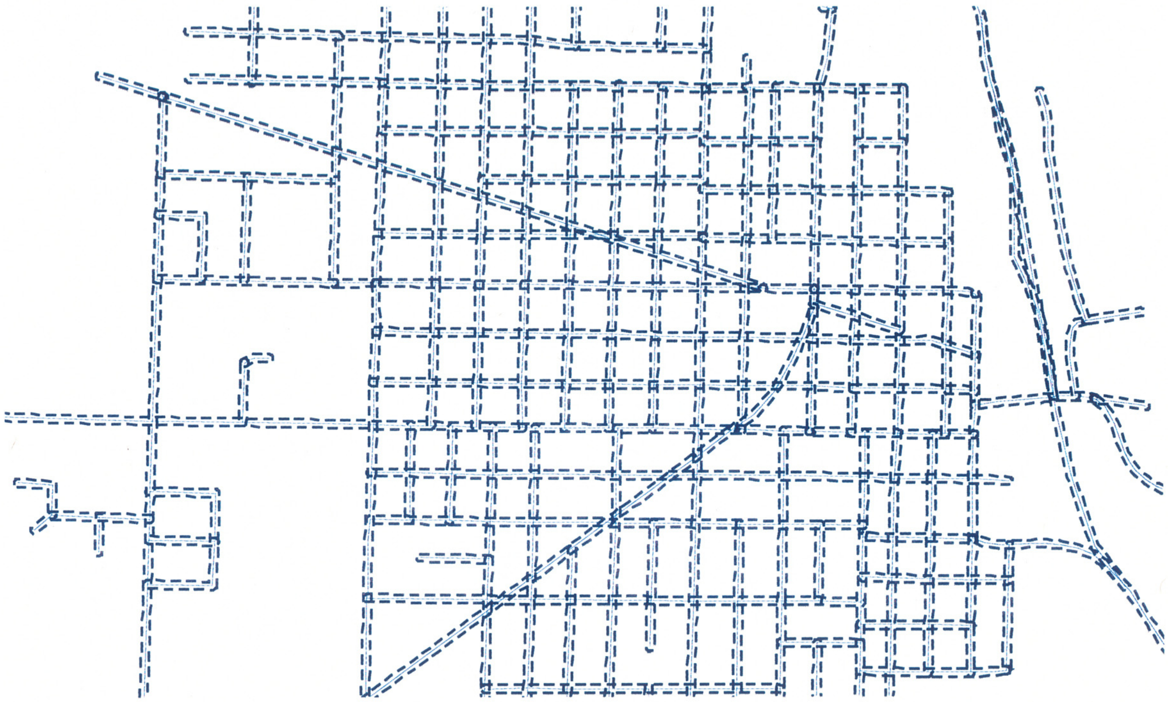


Figure 2.4 Zoomed-In View of Buffer Zone of 10 Meters

The records closest to the subject record are matched first. Once a TIGER ® record is matched with an ITD record it is deleted from the database so that the tool will not have to go through the process of matching it again in the next iteration. The matching process is repeated until the buffer zone is relatively large, with the width in the 101 to 150 meter range. Figures 2.5 and 2.6 show the results from the matching process for 10 and 150-meter buffer zones, respectively.

In these figures dashed lines show TIGER features while ITD features are shown in solid lines.

For this case there were 1283 TIGER ® records at the beginning of the conflation; only 213 remained at the end of the iterations. These remaining records could not be matched with any ITD records. Either one or both of the data sources do not have an accurate representation of the street network in Sandpoint. This phenomenon was observed in other cities as well.

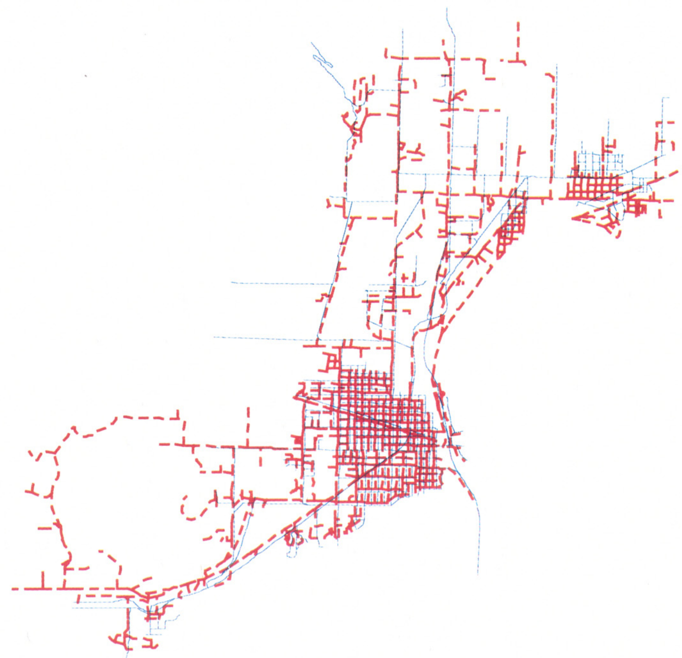


Figure 2.5 Map After 10-Meter Match Iteration

Figure 2.6 Map After 150-Meter Match Iteration

After the matching process is completed the conflation tool has other functions that can be used to make various tables required by the MACS database. The Roadway Names Table has fields for roadway names and address ranges. When the first matching TIGER ® record is found for an ITD road segment the name of the TIGER ® record is assigned to it. During the matching process other TIGER ® records are added to the matched record set for a given ITD record only if the names of the other TIGER ® records match the name of the TIGER ® record that is matched first with the ITD record.

Most of matching proceeds without any intervention from the user. But the user needs to check for errors after every iteration. One type of mismatch occurs when a TIGER ® record within the buffer zone of the ITD record does not have a name. The conflation software adds subsequent TIGER ® records to the matched set only if their names match with the name of the first TIGER ® record matched to the ITD record. When the names do not match, the user needs to manually add the unmatched TIGER ® records to the matched set.

After the name field in the Roadway Name Table is populated, the tool collects address range information from both sides of the matched records and populates the beginning and ending address fields for both the right and left sides of the road. When gathering the address range information the program takes into account the fact that the direction of not all of the TIGER ® chains comprising the matched set for a given ITD record may be the same. When the direction is not the same for all of the chains, care must be taken to ensure that the address range from the correct side of the road is being added.

Conflation of attribute information from TIGER ® with spatial data from ITD is thus achieved through the use of software that was developed for this project. In the sections that follow, the basic functionalities of the conflation tool are discussed in more detail.

2.3 Functions of The Conflation Tool

2.3.1 Buffer zone

The tool draws a buffer zone around each of the ITD segments. This will allow the user to see the number of records (features) of TIGER ® data that are inside the buffer zone. As mentioned before, the user has the ability to specify the width of the buffer zone starting from one meter. Initially, a low buffer zone width should be chosen because this will ensure that during the matching process a lesser number of TIGER ®features will need to be processed. If a bigger width is specified then the matching process involves more number of TIGER ® records which makes the task more complicated and lengthy. The likelihood of more mismatches will also increase.

2.3.2 Match

This is the most crucial step in the whole process. The execution of other commands and the accuracy of the subsequent results directly depend on this "match" function. The match function matches the records and creates a table named "Highways" in the work database showing the information about the records that were matched in that particular iteration. Subsequent operations use the data from this Highways table, the TIGER data, and other data from the "conflat_support_tables" that were created as a part of this project. The conflat_support_tables is an ACCESS database with various unpopulated tables that are used by the program to create the tables required for this project.

Using this command, the ITD segments and TIGER ® segments that are present within the buffer zone are matched. This command searches for the nearest TIGER ® record and assigns it to the ITD record. In subsequent searches the program will seek TIGER ® records with the same name as the one first assigned to the ITD record.

If a TIGER ® segment closest to the ITD record does not have a name, the program will skip this segment and pick the next closest TIGER ® segment. Errors can occur in this process. Wrong sets of TIGER ® segments can be assigned to an ITD record simply because one of the TIGER ® segments happened to be close to the ITD record.

Thus it is very important to check the results manually after every iteration of the matching process. This is necessary because the alignment of both the ITD record and the corresponding TIGER ® segments maybe separated by a distance greater than the width of the user-defined buffer zone.

Once a mismatch is found (generally the names FENAME and ITDNAME are compared to see if they are same or not) the ID of the ITD segment and the GAVPrimaryKey of the TIGER ® segment are noted. Then in the map window the ITD segment is highlighted (by clicking on the corresponding ID in the "Highways" table) and compared with the matched TIGER ® record to see if the two match or not. If a mismatch is found the nearest possible matching ITD record corresponding to the TIGER ® record is sought. Looking for similar names and alignments will help in this step.

If a matching ITD record is found, the TIGER ® record is added as a matching segment to the ITD record. If a match is not found this TIGER ® record is deleted. Care should be taken to adjust the number in the count field of the Highways table when such manual matching is done.

Once the manual process of verifying mismatches is complete, the process is repeated for the next bigger buffer zone. Generally, the matching is done from 1-10 m first, by incrementing the distance by 1 meter between iterations. This is followed by 11-20, 21-30, and 31-40 m range. By the end of the 40 m iteration, most of the TIGER ® records will be matched with the nearest ITD records and the volume of data remaining to be processed will reduce progressively.

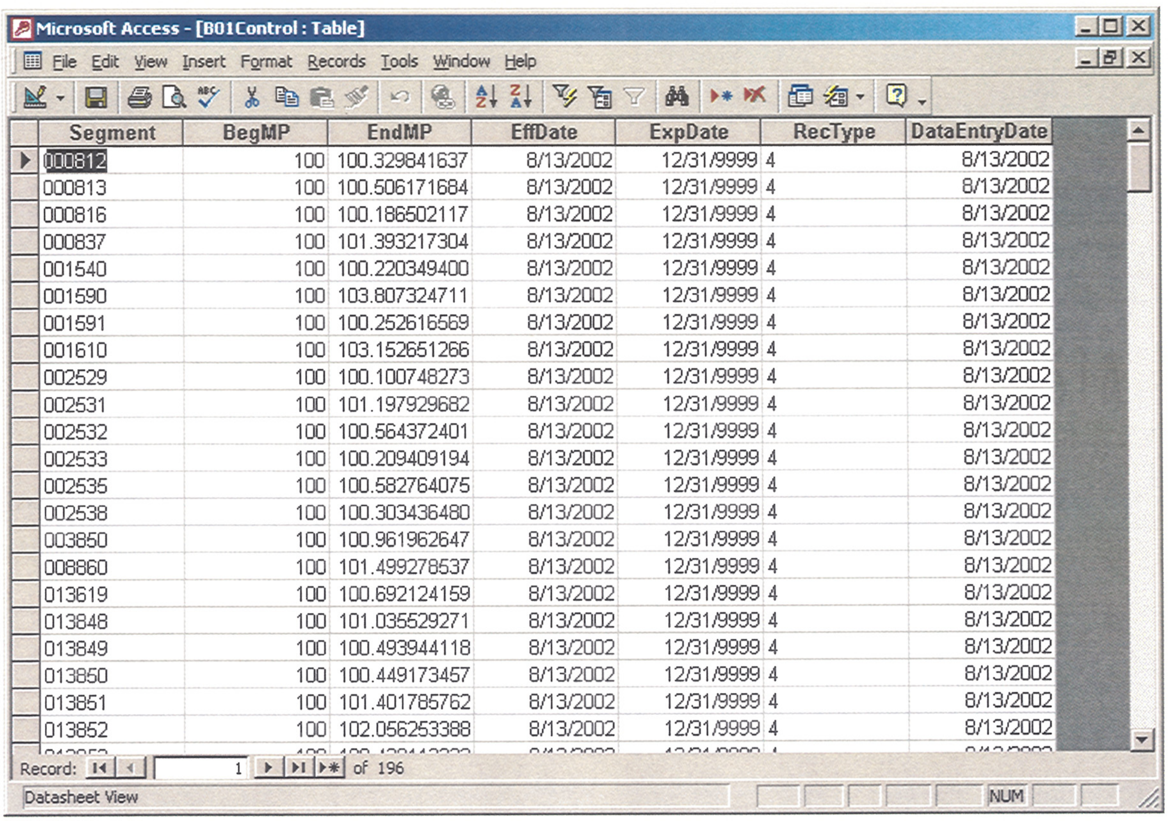
The next matching will be from 41-100 m and finally from 101-150 m. After this iteration and other earlier iterations, mismatches are checked. To repeat, large buffer zones are not used in the earlier phases of matching. If there is a mismatch in a certain iteration, this error continues in subsequent iterations and the errors multiply in the process. So it is important to ensure that this type of errors does not occur at earlier stages since a lot of manual work will be needed later to correct this problem.

2.3.3 Make

This function is used once the "matching" is completed for all the records. "Make" is used to fill in the data for B01, B05, B11 and B16 tables. Of these, B11 and B16 tables will require further manual work, as described later.

2.3.3.1 B01 Control Table

This table contains the information such as segment code, beginning and ending mile points of the road segment and the effective dates generated by the program. An image of this table is shown in Figure 2.7 below.



Segment	BegMP	EndMP	EffDate	ExpDate	RecType	DataEntryDate
000812	100	100.329841637	8/13/2002	12/31/9999	4	8/13/2002
000813	100	100.506171684	8/13/2002	12/31/9999	4	8/13/2002
000816	100	100.186502117	8/13/2002	12/31/9999	4	8/13/2002
000837	100	101.393217304	8/13/2002	12/31/9999	4	8/13/2002
001540	100	100.220349400	8/13/2002	12/31/9999	4	8/13/2002
001590	100	103.807324711	8/13/2002	12/31/9999	4	8/13/2002
001591	100	100.252616569	8/13/2002	12/31/9999	4	8/13/2002
001610	100	103.152651266	8/13/2002	12/31/9999	4	8/13/2002
002529	100	100.100748273	8/13/2002	12/31/9999	4	8/13/2002
002531	100	101.197929682	8/13/2002	12/31/9999	4	8/13/2002
002532	100	100.564372401	8/13/2002	12/31/9999	4	8/13/2002
002533	100	100.209409194	8/13/2002	12/31/9999	4	8/13/2002
002535	100	100.582764075	8/13/2002	12/31/9999	4	8/13/2002
002538	100	100.303436480	8/13/2002	12/31/9999	4	8/13/2002
003850	100	100.961962647	8/13/2002	12/31/9999	4	8/13/2002
008860	100	101.499278537	8/13/2002	12/31/9999	4	8/13/2002
013619	100	100.692124159	8/13/2002	12/31/9999	4	8/13/2002
013848	100	101.035529271	8/13/2002	12/31/9999	4	8/13/2002
013849	100	100.493944118	8/13/2002	12/31/9999	4	8/13/2002
013850	100	100.449173457	8/13/2002	12/31/9999	4	8/13/2002
013851	100	101.401785762	8/13/2002	12/31/9999	4	8/13/2002
013852	100	102.056253388	8/13/2002	12/31/9999	4	8/13/2002

Figure 2.7 B01 Control Table

2.3.3.2 B05 County Limits Table

This table contains information about the counties to which various road segments belong. Details such as the county name and number and boundary codes are listed for each road segment. An image of this table is shown in Figure 2.8 below.

Microsoft Access - [B05CountyLimits : Table]

Segment	BegMP	EndMP	EffDate	ExpDate	CountyNo	CountyName	BoundaryCode	DataEntryDate
000812	100	100.329841637	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
000813	100	100.506171684	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
000816	100	100.186502117	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
000837	100	101.393217304	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
001540	100	100.220349400	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
001590	100	103.807324711	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
001591	100	100.252616569	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
001610	100	103.152651266	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
002529	100	100.100748273	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
002531	100	101.197929682	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
002532	100	100.564372401	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
002533	100	100.209409194	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
002535	100	100.582764075	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
002538	100	100.303436480	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
003850	100	100.961962647	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
008860	100	101.499278537	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
013619	100	100.692124159	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
013848	100	101.035529271	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
013849	100	100.493944118	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
013850	100	100.449173457	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
013851	100	101.401785762	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
013852	100	102.056253388	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
013853	100	100.439113232	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
013854	100	100.19865226	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
015909	100	101.755222394	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002
016003	100	100.343958822	8/13/2002	12/31/9999	16017	BONNER	1	8/13/2002

Record: 1 of 196

Datasheet View

Figure 2.8 B05 County Limits Table

2.3.3.3 B11 Jurisdiction Table

The Jurisdiction table contains details such as jurisdiction codes on both left and right sides of the road segments. The tool is designed in such a way that any feature that is contained entirely within the borders of a city limit is considered to be within the city limit and hence the jurisdiction code on the left side of the road segment (denoted by "JurisdictionLt" in the table) as well as the code on the right side of the segment (denoted by "JurisdictionRt" in the table) are both taken to be the value entered in the JurisdictionIn column in the city_FIPS table in the conflat_support_tables database.

Even if a very small portion of a segment falls outside the city limit or the segment is ending/beginning at the border of a city limit, the segment is considered to be outside the jurisdiction. The codes for both left and right side of such a segment are chosen from the JurisdictionOut code in the city_FIPS table. In some cases, a part of road segment may lie inside while the rest is outside. In this case also, both codes (Lt and Rt) are taken from the JurisdictionOut field. And in some other cases, a road segment may be lying exactly on the border of the city limit, collinear with it completely. In such cases, there will be two different codes for the jurisdiction left and right sides of the segments.

Microsoft Access - [B11Jurisdiction : Table]

Segment	BegMP	EndMP	EffDate	ExpDate	JurisdictionLt	JurisdictionRt	Gated	DataEntryDate	Done
000812	100	100.329841637	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
000813	100	100.506171684	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
000816	100	100.186502117	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
000837	100	101.393217304	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
001540	100	100.220349400	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
001590	100	103.807324711	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
001591	100	100.252616569	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
001610	100	103.152651266	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
002529	100	100.100748273	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
002531	100	101.197929682	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
002532	100	100.564372401	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
002533	100	100.209409194	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
002535	100	100.582764075	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
002538	100	100.303436480	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
003850	100	100.961962647	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
008860	100	101.499278537	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
013619	100	100.692124159	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
013848	100	101.035529271	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
013849	100	100.493944118	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
013850	100	100.449173457	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
013851	100	101.401785762	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
013852	100	102.056253388	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
013853	100	100.439113232	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
013854	100	100.19865226	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
015909	100	101.755222394	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
016003	100	100.343958822	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
016005	100	100.233544316	8/13/2002	12/31/9999	3200	3200	No	8/13/2002	-1
022101	100	101.172095055	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
022180	100	100.770885196	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
022181	100	100.529994026	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
022182	100	100.243901418	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
022183	100	100.154346302	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1
022184	100	100.372349194	8/13/2002	12/31/9999	90	90	No	8/13/2002	-1

Record: 1 of 196

Datasheet View

Figure 2.9 B11 Jurisdiction Table

The Conflation tool cannot fill those codes automatically. They must be entered manually. Therefore it is necessary to check for such segments in the map once the "make" is done. Looking simultaneously at the map window of the geoworkspace and the data window of the B16 table/Highways table and entering the proper codes in the table the manual entry of jurisdiction codes can be completed. An image of the Jurisdiction table in shown above in Figure 2.9.

2.3.3.4 B16 RoadwayNames_highways table

The B16 Roadway Names table contains data obtained from the final "Highways" table (which is created after many iterations of matching process) and the TIGER ® data. This table also contains manually entered data such as the sequence numbers. The sequence numbers are required to keep track of various segments of a road with the same name. The numbering is assigned on the basis of a predefined convention. It is assumed that all roads start either from the west or south. The first segment is given a sequence number of 50 and the rest are given 60,70,80 and so on.

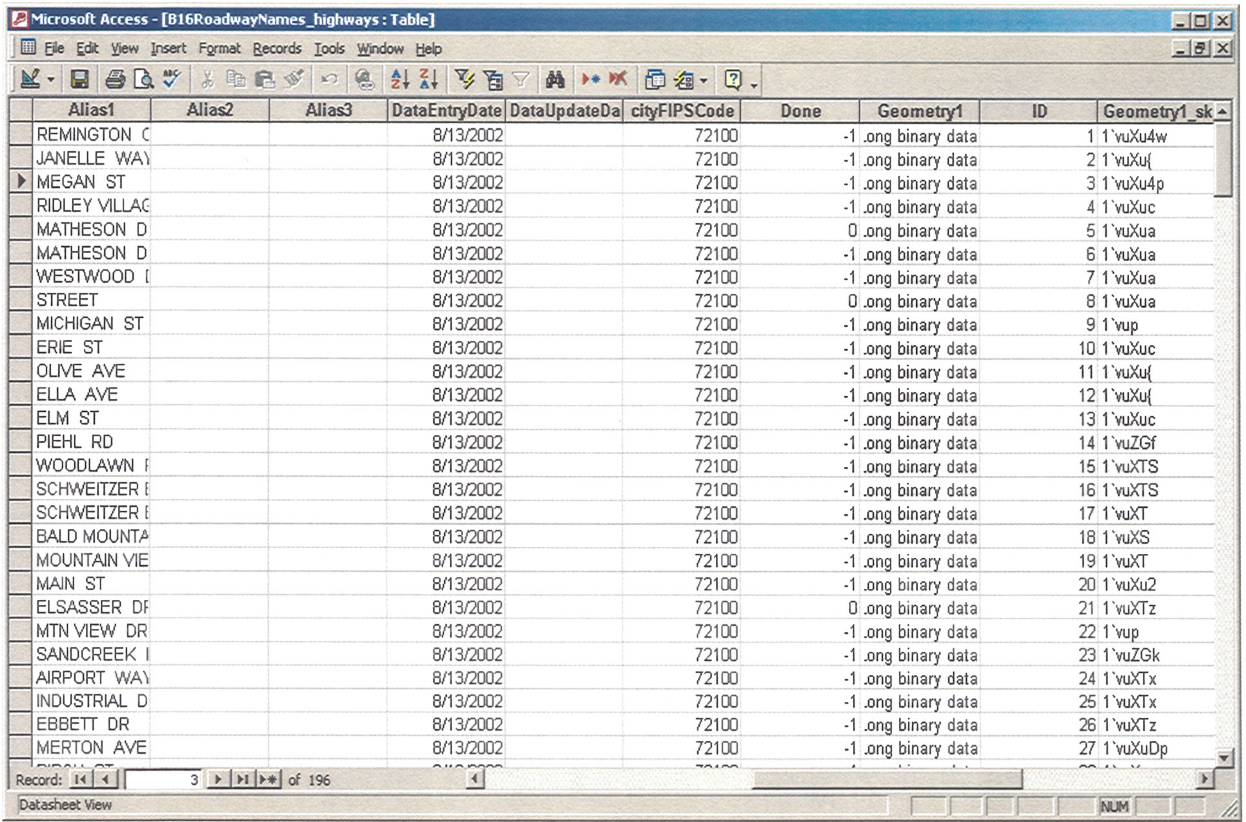
Microsoft Access - [B16RoadwayNames_highways : Table]

File Edit View Insert Format Records Tools Window Help

Segment	BegMP	EndMP	EffDate	ExpDate	Prefix	Name	Suffix1	Suffix2	SeqNo	BegAddressLt	BegAddressRt	EndAddressLt	EndAddressRt
022227	100	100.054	3/13/2002	2/31/9999		Remingtor		Ct		301	300	417	418
022224	100	100.352	3/13/2002	2/31/9999		Janelle		Way		1701	112	2099	2098
022225	100	100.104	3/13/2002	2/31/9999		Megan		Ln		1801	1800	2099	2098
022223	100	100.402	3/13/2002	2/31/9999		Ridley Vill:		Rd		801	800	1399	1816
022221	100	100.385	3/13/2002	2/31/9999									
022220	100	100.302	3/13/2002	2/31/9999		Mathison		Dr		1301	1300	1699	1698
062001	100	100.213	3/13/2002	2/31/9999		Westwood		Dr		101	100	1899	1898
022222	100	100.086	3/13/2002	2/31/9999									
022245	100	100.770	3/13/2002	2/31/9999	W	Michigan		St		401	400	1399	1398
022243	100	100.461	3/13/2002	2/31/9999		Erie		St		501	500	1199	1198
022241	100	100.68	3/13/2002	2/31/9999	S	Olive		Ave		201	116	699	698
022240	100	100.68	3/13/2002	2/31/9999	S	Ella		Ave	50	201	200	699	614
062002	100	100.129	3/13/2002	2/31/9999	W	Elm		St	50	1103	1100	1299	1298
022201	100	100.256	3/13/2002	2/31/9999		Piehl		Rd		1	2	699	698
022199	100	101.050	3/13/2002	2/31/9999		Woodland		Dr	50	169	168	999	4880
002535	100	100.583	3/13/2002	2/31/9999		Schweitze		Rd		1	2	699	698
022101	100	101.172	3/13/2002	2/31/9999									
003850	100	100.962	3/13/2002	2/31/9999		Baldy Mou		Rd	50				
022195	100	100.987	3/13/2002	2/31/9999		Mountain \		Dr		21	20	37	36
002529	100	100.101	3/13/2002	2/31/9999	W	Main		St	50	1901	2006	2099	2098
022293	100	100.257	3/13/2002	2/31/9999									
062008	100	100.312	3/13/2002	2/31/9999		Mt View		Rd		401	400	899	898
022294	100	100.288	3/13/2002	2/31/9999		Sandcreek		Ln		401	600	699	698
022290	100	100.300	3/13/2002	2/31/9999		Airport		Way		101	100	1199	1198
022289	100	100.312	3/13/2002	2/31/9999		Industrial		Dr		1701	1700	2099	2098
062009	100	100.24	3/13/2002	2/31/9999		Ebbett		Way		501	448	999	998
022239	100	100.096	3/13/2002	2/31/9999	S	Merton		Ave			702		708

Record: 3 of 196

Datasheet View NUM



Microsoft Access - [B16RoadwayNames_highways : Table]

Alias1	Alias2	Alias3	DataEntryDate	DataUpdateDa	cityFIPSCode	Done	Geometry1	ID	Geometry1_sk
REMINGTON C			8/13/2002		72100	-1	.ong binary data	1	1`vuXu4w
JANELLE WAY			8/13/2002		72100	-1	.ong binary data	2	1`vuXu{
MEGAN ST			8/13/2002		72100	-1	.ong binary data	3	1`vuXu4p
RIDLEY VILLAG			8/13/2002		72100	-1	.ong binary data	4	1`vuXuc
MATHESON D			8/13/2002		72100	0	.ong binary data	5	1`vuXua
MATHESON D			8/13/2002		72100	-1	.ong binary data	6	1`vuXua
WESTWOOD [			8/13/2002		72100	-1	.ong binary data	7	1`vuXua
STREET			8/13/2002		72100	0	.ong binary data	8	1`vuXua
MICHIGAN ST			8/13/2002		72100	-1	.ong binary data	9	1`vup
ERIE ST			8/13/2002		72100	-1	.ong binary data	10	1`vuXuc
OLIVE AVE			8/13/2002		72100	-1	.ong binary data	11	1`vuXu{
ELLA AVE			8/13/2002		72100	-1	.ong binary data	12	1`vuXu{
ELM ST			8/13/2002		72100	-1	.ong binary data	13	1`vuXuc
PIEHL RD			8/13/2002		72100	-1	.ong binary data	14	1`vuZGf
WOODLAWN I			8/13/2002		72100	-1	.ong binary data	15	1`vuXTS
SCHWEITZER I			8/13/2002		72100	-1	.ong binary data	16	1`vuXTS
SCHWEITZER I			8/13/2002		72100	-1	.ong binary data	17	1`vuXT
BALD MOUNTA			8/13/2002		72100	-1	.ong binary data	18	1`vuXS
MOUNTAIN VIE			8/13/2002		72100	-1	.ong binary data	19	1`vuXT
MAIN ST			8/13/2002		72100	-1	.ong binary data	20	1`vuXu2
ELSASSER DI			8/13/2002		72100	0	.ong binary data	21	1`vuXTz
MTN VIEW DR			8/13/2002		72100	-1	.ong binary data	22	1`vup
SANDCREEK I			8/13/2002		72100	-1	.ong binary data	23	1`vuZGk
AIRPORT WAY			8/13/2002		72100	-1	.ong binary data	24	1`vuXTx
INDUSTRIAL D			8/13/2002		72100	-1	.ong binary data	25	1`vuXTx
EBBETT DR			8/13/2002		72100	-1	.ong binary data	26	1`vuXTz
MERTON AVE			8/13/2002		72100	-1	.ong binary data	27	1`vuXuDp

Record: 3 of 196

Datasheet View

Figure 2.10 B16Roadway Names Table

The other data contained in the B16 table are beginning and ending addresses. The program assumes that all left addresses have odd values and right addresses have the even values. Other fields in this table include names, beginning and ending mile points, prefixes, suffixes, aliases, cityfips codes, and the geometry information. An example of the B16 table is shown in Figure 2.10.

2.3.4 C01Make

This function is used to make the C01MP Descriptions table. Before using this command, the "intersections" table needs to be created and stored in the work database. The Spatial Intersection tool from GeoMedia is used to do that. The dialog box for this tool is shown in Figure 2.11 below. Using this tool, the "intersections" table can be made and intersections can be shown on the map window as a feature class.

Both the input and output features are selected as B16RoadwayNames_highways in the work database. Output features are selected as static, editable feature class. The connection name would be that of the work database and the Feature class would be "intersection".

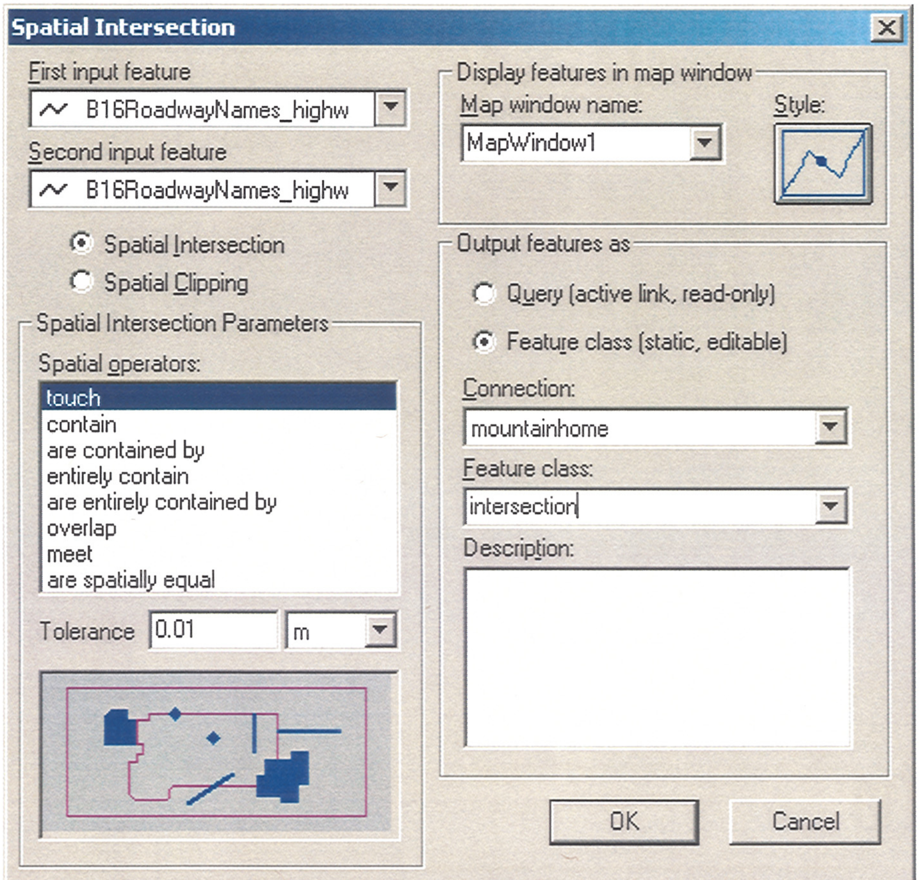


Figure 2.11 Spatial Intersection Tool Dialogue Box

Once the "intersections" table is made, the "MakeC01" command can be used. This command creates the "C1MPDescriptions" table in the work database. This table uses the data from the "intersections" table created by the Spatial Intersection tool. It contains data related to the intersections of various segments, the location of the intersection measured from the start of the segment, and the corresponding mile point description codes as specified by the general rules in MACS / ROSE access guide. The guide has specific codes for a 2-leg, 3-leg, and 4- or more leg intersections. In the case of a 3-leg intersection, a sub-classification is also given to distinguish among 3-leg left, 3-leg right and 3-leg left & right intersection. An image of a C01MP table created by the Conflation tool is shown in Figure 2.12 below. The figure has two parts and shows all the fields in the table.

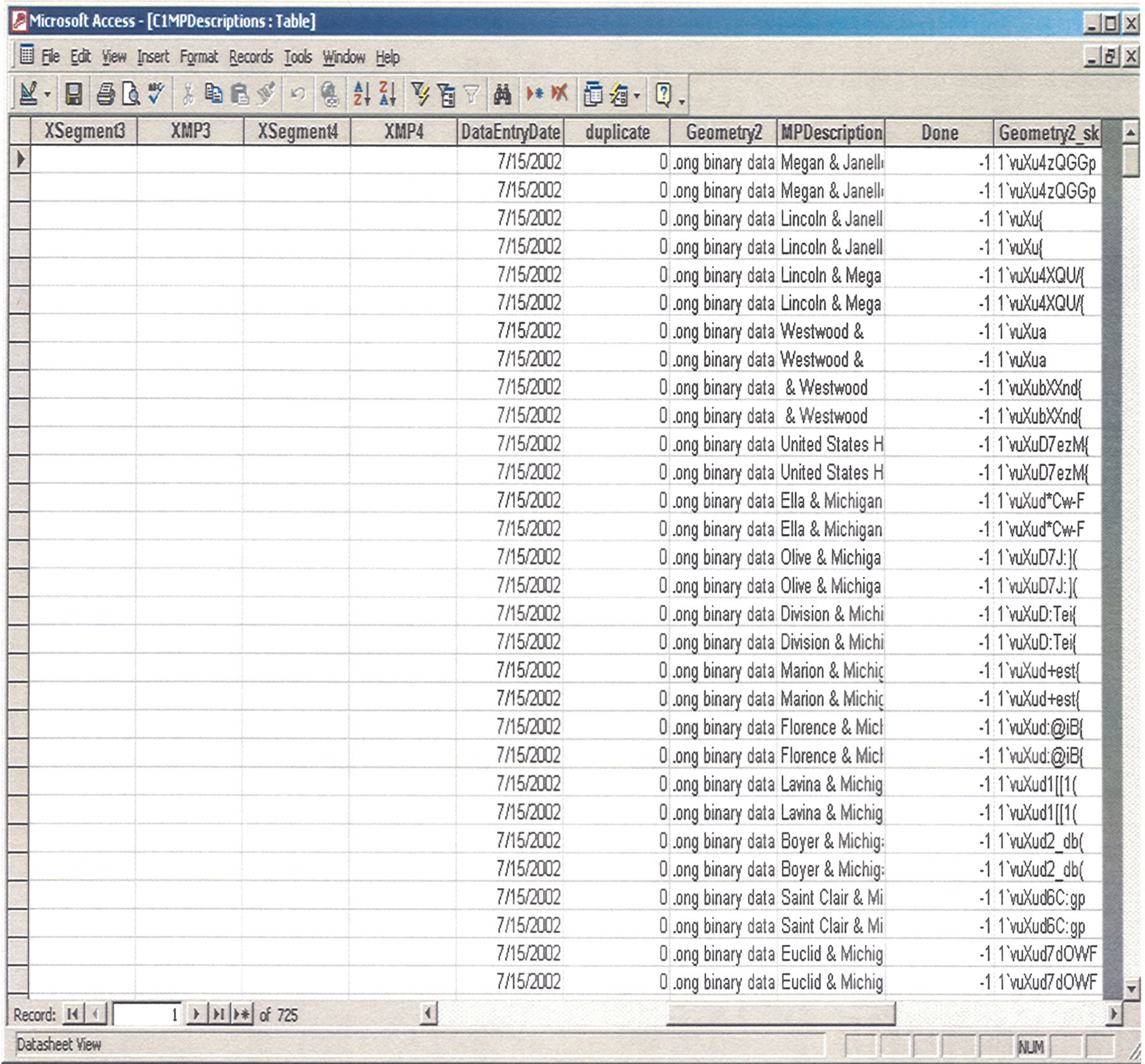
Microsoft Access - [C1MPDescriptions : Table]

File Edit View Insert Format Records Tools Window Help

ID	Segment	MP	EffDate	ExpDate	MPDescCode	XSegment1	XMP1	XSegment2	XMP2
1	022224	100.171723908	4/24/2002	12/31/9999	29	022225	100.102932378		
2	022225	100.102932378	4/24/2002	12/31/9999	31	022224	100.171723908		
3	022224	100.350392531	4/24/2002	12/31/9999	31	013848	100.421315695		
4	013848	100.421315695	4/24/2002	12/31/9999	30	022224	100.350392531		
5	022225	100	4/24/2002	12/31/9999	31	013848	100.345649989		
6	013848	100.345649989	4/24/2002	12/31/9999	30	022225	100		
7	022221	100.307332491	4/24/2002	12/31/9999	30	062001	100.211263503		
8	062001	100.211263503	4/24/2002	12/31/9999	31	022221	100.307332491		
9	062001	100.130486248	4/24/2002	12/31/9999	30	022222	100.085877673		
10	022222	100.085877673	4/24/2002	12/31/9999	31	062001	100.130486248		
11	022245	100.210510408	4/24/2002	12/31/9999	32	001590	103.239970273		
12	001590	103.239970273	4/24/2002	12/31/9999	32	022245	100.210510408		
13	022245	100.2553338	4/24/2002	12/31/9999	32	022240	100.530100408		
14	022240	100.530100408	4/24/2002	12/31/9999	32	022245	100.2553338		
15	022245	100.186921109	4/24/2002	12/31/9999	32	022241	100.409681499		
16	022241	100.409681499	4/24/2002	12/31/9999	32	022245	100.186921109		
17	022245	100	4/24/2002	12/31/9999	31	008860	100.159387157		
18	008860	100.159387157	4/24/2002	12/31/9999	30	022245	100		
19	022245	100.317651178	4/24/2002	12/31/9999	32	022242	100.270690041		
20	022242	100.270690041	4/24/2002	12/31/9999	32	022245	100.317651178		
21	022245	100.382620636	4/24/2002	12/31/9999	32	022244	100.271894312		
22	022244	100.271894312	4/24/2002	12/31/9999	32	022245	100.382620636		
23	022245	100.442633155	4/24/2002	12/31/9999	32	022247	100.073021271		
24	022247	100.073021271	4/24/2002	12/31/9999	32	022245	100.442633155		
25	022245	100.508733996	4/24/2002	12/31/9999	32	062036	100.269847268		
26	062036	100.269847268	4/24/2002	12/31/9999	32	022245	100.508733996		
27	022245	100.579446253	4/24/2002	12/31/9999	32	022248	100.207084595		
28	022248	100.207084595	4/24/2002	12/31/9999	32	022245	100.579446253		
29	022245	100.647486463	4/24/2002	12/31/9999	32	013850	100.183585979		
30	013850	100.183585979	4/24/2002	12/31/9999	32	022245	100.647486463		

Record: 1 of 725

Datasheet View NUM



XSegment3	XMP3	XSegment4	XMP4	DataEntryDate	duplicate	Geometry2	MPDescription	Done	Geometry2_sk
				7/15/2002	0	.ong binary data	Megan & Janell	-1	1`vuXu4zQGGp
				7/15/2002	0	.ong binary data	Megan & Janell	-1	1`vuXu4zQGGp
				7/15/2002	0	.ong binary data	Lincoln & Janell	-1	1`vuXu{
				7/15/2002	0	.ong binary data	Lincoln & Janell	-1	1`vuXu{
				7/15/2002	0	.ong binary data	Lincoln & Mega	-1	1`vuXu4XQU/{
				7/15/2002	0	.ong binary data	Lincoln & Mega	-1	1`vuXu4XQU/{
				7/15/2002	0	.ong binary data	Westwood &	-1	1`vuXua
				7/15/2002	0	.ong binary data	Westwood &	-1	1`vuXua
				7/15/2002	0	.ong binary data	& Westwood	-1	1`vuXubXXnd{
				7/15/2002	0	.ong binary data	& Westwood	-1	1`vuXubXXnd{
				7/15/2002	0	.ong binary data	United States H	-1	1`vuXuD7ezM{
				7/15/2002	0	.ong binary data	United States H	-1	1`vuXuD7ezM{
				7/15/2002	0	.ong binary data	Ella & Michigan	-1	1`vuXud*Cw-F
				7/15/2002	0	.ong binary data	Ella & Michigan	-1	1`vuXud*Cw-F
				7/15/2002	0	.ong binary data	Olive & Michiga	-1	1`vuXuD7J:](
				7/15/2002	0	.ong binary data	Olive & Michiga	-1	1`vuXuD7J:](
				7/15/2002	0	.ong binary data	Division & Michi	-1	1`vuXuD:Tei{
				7/15/2002	0	.ong binary data	Division & Michi	-1	1`vuXuD:Tei{
				7/15/2002	0	.ong binary data	Marion & Michig	-1	1`vuXud+est{
				7/15/2002	0	.ong binary data	Marion & Michig	-1	1`vuXud+est{
				7/15/2002	0	.ong binary data	Florence & Mich	-1	1`vuXud:@iB{
				7/15/2002	0	.ong binary data	Florence & Mich	-1	1`vuXud:@iB{
				7/15/2002	0	.ong binary data	Lavina & Michig	-1	1`vuXud1[[1(
				7/15/2002	0	.ong binary data	Lavina & Michig	-1	1`vuXud1[[1(
				7/15/2002	0	.ong binary data	Boyer & Michig	-1	1`vuXud2_db(
				7/15/2002	0	.ong binary data	Boyer & Michig	-1	1`vuXud2_db(
				7/15/2002	0	.ong binary data	Saint Clair & Mi	-1	1`vuXud6C:gp
				7/15/2002	0	.ong binary data	Saint Clair & Mi	-1	1`vuXud6C:gp
				7/15/2002	0	.ong binary data	Euclid & Michig	-1	1`vuXud7dOWF
				7/15/2002	0	.ong binary data	Euclid & Michig	-1	1`vuXud7dOWF

Figure 2.12 C1MP Descriptions Table

3. PERFORMANCE

The performance of the conflation tool can be evaluated in a number of ways. First, the total number of times mismatches, that occur in the matching process as the buffer zone is widened from 1 to 150 m, is recorded. The number of records that is imported initially is also recorded. Table 3.1 below lists these data for seven of the eight cities processed as a part of this research. Information for the city of Mountain Home is not shown, as this data was not collected when this city was processed.

The fourth column in the table shows the number of instances of mismatches as the processing progresses. Each mismatch corresponds to an ITD record. The number listed is not the total number of unique ITD records that were involved in a mismatch. It just indicates the total number of instances of mismatch that were observed during the entire process. For example, the city of Sandpoint had 106 instances of mismatches. This does not mean that there were 106 separate ITD records that had mismatches. It just means that mismatches were observed 106 times and one ITD record could have been involved in a mismatch more than once during the various iterations.

The number of manual matches that were performed is shown in the fifth column of the table. This number indicates the actual number of unique ITD records for which manual matching had to be performed. In the case of Sandpoint there were 83 such records. Also, the two groups of records, namely the two sets of 83 and 106 records in the case of Sandpoint, need not be mutually exclusive. In other words, records that had mismatches in the automated processing stage, could also have been involved in the manual matching stage.

As mismatches and manual matches also involve TIGER ® records, the number of instances of these events is directly correlated with the number of TIGER ® records present in the database. The total number of mismatches and matches as a percentage of the TIGER ® records for a city has been computed and is shown in the table.

Table 3.1 The Number Of Mismatches

City	Number of ITD Records	Number of TIGER Records	Instances of Mismatches (1 – 150 m)	Manual Matches After 150 m	Mismatches & Matches	Mismatches & Matches As a % of TIGER Records
Caldwell	619	2790	98	51	149	5.3%
Coeur d' Alene	816	4468	135	167	302	6.8%
Lewiston	604	2701	245	106	351	13.0%
Moscow	345	1424	46	34	80	5.6%
Nampa	1013	4907	259	108	367	7.5%
Post Falls	392	2229	111	42	153	6.9%
Sandpoint	196	1283	106	83	189	14.7%

As can be seen from the table this percentage ranges from a low of 5.3% for Caldwell to 14.7% for Sandpoint. The value of this measure is less than 10% for five of seven cities listed in the table. The aggregate percentage of total number of mismatches and manual matches for all seven cities in relation to the total number of TIGER ® records is 8%.

Table 3.2 shows the percentage of TIGER ® records matched after the automated processing as well as the manual processing is completed. This number after the completion of all processing ranged from about 74% for Post Falls to over 93% for Nampa.

The final table, Table 3.3, shows the distribution of mismatches during the automated processing as the processing proceeds from a narrow buffer zone to progressively larger zones.

Table 3.2 Performance Based On The Percentage Of Tiger Records Matched

City	Number of ITD Records	Number of TIGER ® Records Before Matching	Number of TIGER ® Records After Matching up to 150 Meters	Percentage of TIGER ® Records Matched	Number of TIGER ® Records After Manual Matching	Final Percentage of TIGER ® Records Matched
Caldwell	619	2790	569	79.61	467	83.26
Coeur d' Alene	816	4468	1125	74.82	815	81.76
Lewiston	604	2701	429	84.12	232	91.41
Moscow	345	1424	316	77.81	217	84.76
Nampa	1013	4907	704	85.65	322	93.44
Post Falls	392	2229	667	70.08	586	73.71
Sandpoint	196	1283	419	67.34	213	83.40

Table 3.3 Distribution Of Mismatches

City	Mismatches At The End Of Iteration															Total
	1m	2m	3m	4m	5m	6m	7m	8m	9m	10m	11-20m	21-30m	31-40m	41-100m	101-150m	
Caldwell	0	N/A	2	N/A	2	3	1	0	2	1	19	14	8	24	22	98
Coeur d' Alene	0	0	0	2	0	2	N/A	N/A	N/A	N/A	33	24	14	27	33	135
Lewiston	0	0	4	0	1	0	1	5	1	3	26	44	46	85	19	234
Moscow	0	1	0	0	0	1	0	0	0	1	11	7	6	11	8	46
Nampa	0	1	3	2	1	4	1	2	3	4	45	29	30	88	46	259
Post Falls	0	1	1	0	0	0	2	N/A	N/A	N/A	19	25	18	N/A	45	111
Sandpoint	0	0	0	1	0	0	0	1	0	0	6	15	21	39	23	106

N/A: Data not available

4. TOOL VALIDATION

The eight urban areas processed in this project are: Caldwell, Coeur d'Alene, Lewiston, Moscow, Mountain Home, Nampa, Sandpoint and Post Falls. Of these eight, three are in Southwest Idaho, close to the capital city of Boise. Of these three cities, Nampa is the closest and also the biggest. The validation of the tool output was done first in Nampa. The results from this validation process are reported here.

4.1 Sample Selection

There are 1013 ITD records (road segments) representing the road network within the urban area limits of the city of Nampa. The corresponding number of TIGER ® records is 4907. After the matching process was completed, 322 TIGER ® remained. Thus not all TIGER ® records could be matched with an ITD record. Likewise, not all ITD records got matched with a TIGER ® record.

4.1.1 Sample Size

Fifteen ITD records were selected for the validation. This constitutes about 1.5% of the total number of records. The sampling was not random, but it was not selected by BSU, the developers of the tool. Personnel from ITD selected the sample.

As explained by these personnel, an even distribution in terms of location, type of street, and adjoining land use were the objectives in the selection of the sample. Table1 lists the fifteen road segments along with other pertinent data.

The first field in the table is a six-digit code for the segment. This code is used along with reference mile points and dates to form a unique key by the MACS system. The next column heading is the name of road segment. The name used is from the FENAME field in the TIGER ® database. The length of the segment is noted next. The next column shows the separation of the edge of the buffer zone from the road feature, at the time the first matching TIGER ® record was found. Finally, the count is the number of TIGER ® records matched with the ITD record.

Table 4.1 List Of Validation Sample

Segment #	Name	Length(miles)	Distance (meters)	Count
061430	1st Street North Extended	1.6700	1	35
013347	4th St	1.5131	6	21
013346	5th St	1.2774	16	18
013287	10th Ave S	1.6252	5	21
013165	W Blaine Ave	1.2049	13	17
004563	Chicago St	0.8340	13	12
061448	Greenhurst Rd	2.9833	50	42
004720	E Greenhurst Rd	1.0012	28	24
002550	Greenhurst Rd	0.9966	101	8
061447	Greenhurst (No FENAME)	0.6818	–	0
004553	E Rossevelt Ave	0.8324	36	22
000333	Sugar St	1.5706	5	23
028820	Sunflower Dr	0.5341	6	9
012243	Ventura Dr	0.5779	14	8
013263	Wildflower Dr	0.4886	52	9

– no matching TIGER® record found

4.2 Validation of Methodology

The validation was performed by visiting the road location and noting the name of the segment as well as the address on both sides of the road. Validation for the 15 road segments was completed in one day.

The problems observed were of three types:

- Missing/inadequate information in ITD' s database
- Missing/inadequate information in TIGER ®
- Problems the software was not designed to handle

There are 12 different roads represented in the 15 records in Table 4.1. Most of the problems observed in the field verification were common among a few or all of the 12 roads. Presented below is a description of the observations made for each of the roads listed in the table. In some cases the description for some of the roads has been combined, when the problems encountered in them were similar.

4.3. Field Observations

1st Street North Extended

As can be seen from the table, 35 TIGER ® records were matched with this road segment, and the matching process started at a buffer distance of 1m.

The major problem with this street is that in the ITD database it is shown as one record while in fact there are two parts of this road: 1st Street North and 1st Street North Extended. As described previously the conflation tool starts its processing by drawing buffer zones around each ITD record and grabbing TIGER ® records that fall within the buffer zone.

The name used in the Roadway Names Table in the MACS database is the FENAME from TIGER ®. The software starts the processing for each road either from western or the southern ends depending on whether the orientation of the road is east-west or north-south. The processing for this road started from the western end and its name was obtained from the FENAME field from the TIGER ® record that was at that end. Proceeding along this road to the eastern side, it was observed that the name of the TIGER ® records close to the road changed to 1st Street North. But this change was not reflected in the name field of the Roadway Names Table because the software is designed not to change names after it is first assigned.

The problem described occurred because ITD' s database was not current. Because of this discrepancy the address range shown for this road is incorrect.

4th and 5th Streets

These streets are similar and pass through downtown Nampa. No problems were observed in the name of the streets. Address ranges were similar to what was observed in the field. The beginning and ending addresses were not exactly the same as observed in the field. For example on 5th Street according to the database the beginning address on the left and right should have been 301 and 300, respectively. But in the field these numbers were not observed. The starting addresses on the left and right was 311 and 312, respectively.

Another general observation about the actual beginning of the addressing needs to be made here. The linear referencing for the roads starts from the western end for an east-west road and southern end for a north-south road. In other words, the beginning mile point is always on the western or the southern end of the road. Though it would be nice to have a similar convention for the address, it cannot always be assumed that the beginning address always starts on the western or southern end. Thus it is not clear from the Roadway Names Table, at which end the addressing starts. What the database table has is just the range of addresses and it has no information about which end the addressing begins.

10th Ave S

The field verification revealed that this road also had an extension, named 10th Ave S Ext. This fact was not reflected in either of the two database sources. Other than this missing information, the address range in the database was found to be similar to the range observed in the field.

W. Blaine Ave

No major discrepancy was observed for this road. One minor gap in TIGER ® was the missing prefix, W, in the name of the road. The other was related to the magnitude of the ending addresses on the left and right. According to the TIGER ® database these numbers should have been 1399 and 1398. The actual address values observed were 1723 and 1512.

Chicago St, Ventura Dr, Wildflower Dr and Sunflower Dr

The database information for spatial as well as attribute data for these four roads matched with the field observation. The only discrepancy was that the values for the beginning and ending addresses in the field did not match exactly with that of the database.

Greenhurst Rd

Greenhurst Rd is an arterial road that runs east west along the southern part of town. ITD represents this road as four spatial features. TIGER ® road segments exist alongside three of these segments. For the fourth ITD segment, with segment code 061447, no TIGER ® segment exists.

Other than this fact, data from the two sources do not deviate from each other, except for the fact that the values for beginning and ending address do not match exactly between the database source and the field.

E Roosevelt Rd

An error in the ITD spatial data was observed on this road. A portion of the road on the western end of this east-west road is part of a different street, 14th Street S. No other gaps in the data were observed. The conflation tool correctly assigned the name and address range information to the ITD spatial feature.

Sugar St

Both ITD and TIGER ® had errors in their data for this street. In the field this street was found to have two components, North and South. But the ITD database represents these streets as one street. As a result, the address range information for this street is not correct.

The error in the TIGER ® database is that the FETYPE is shown as Ave while it should have been St.

4.4 Validation Summary

Of the 15 segments examined in the field, ten (or 67%) were found to have no database errors. One segment had erroneous ITD spatial data; ITD did not have information about an extended section of an existing road. Two segments had gaps in the TIGER ® database. One of those two was related to a missing prefix and in the other no road segments close to an existing road existed in the TIGER ® database. The two remaining segments had erroneous information from both ITD and TIGER ® databases.

The conflation tool, the object of the validation exercise, was found to function as designed. For each ITD road segment, the tool identified TIGER ® segments that matched with it, collected attribute information from TIGER ® and saved this information in a database. The tool did not identify, however, the correct road segment end for the beginning of the address range. But the tool was not designed with this functionality.

5. CONCLUSIONS

A software tool was designed to merge spatial and attribute information from two different sources to create the transportation framework layer for urban areas in Idaho. Spatial data used was obtained from ITD. The attribute data was from the TIGER ® database.

Field verification revealed that the tool works as designed. Errors were found, however, in the input data. It was found that ITD needed to update its urban area maps and TIGER ® also had some discrepancies.

ITD now has a tool to rapidly generate FGDC-compliant base maps for urban areas in Idaho. As the quantity of the input data improves, the utility of this tool will be enhanced.

Future Work

In its current state the tool is capable of assigning road name and address range attribute data to spatial features representing the road network. The tool cannot, however, identify the road segment end where the addressing begins. This information will be critical in emergency situations, especially for longer road segments. It is recommended that the conflation tool described in this report be improved to incorporate this functionality.

APPENDIX A

Follow the steps shown below to install the conflation tool into the GeoMedia software:

1. Copy match1.dll and match1.ini into the GeoMedia program directory.

2. Launch the DOS window (command prompt), enter the directory that contains GeoMedia.

3. In the command line, enter the following command:

   InstallUsrCmd/prod "GeoMedia professional" match1.dll match1.ini

4. Or if the GeoMedia version is not professional, enter:

   InstallUsrCmd/prod "GeoMedia" match1.dll match1.ini

5. If the install is successful, start GeoMedia and go to the Tools menu, and then to the Customize command. Then when the menu tab, "Custom" in "Categories" list box is chosen, the "Conflate TIGER data into ITD table" in Commands list box should be seen. Choose it and add into anywhere you want in the menu.

The tool can then be accessed from the menu where it was added.

Limits:

This tool was developed to conflate TIGER data into the ITD database. The goal is to create the B16 table and other tables whose table structures were specified. The precise table was considered to be "Highways" in the ITD database. The application can also import MGE data into GeoMedia and combine relative tables into one combined table that has the same function as "Highways".

APPENDIX B

To use this tool to create the B16, B01, B05, B11, and C01 tables, the work database and Geoworkspace need to be set up.

1. Open the "conflat_support_tables" database; fill the city_FIPS table (the names of city, their city FIPS, country code, jurisdictions, and the segment codes which can be used to assign to new records.)
2. Create one new workspace in GeoMedia, connect it to the TIGER database, the ITD database that contains the "Highways" table or the ITD MGE database that contains the required tables, the databases which contain the urban boundary information and the cityLimit information, and the "conflat_support_tables" database which contains the structure of some of required tables and the city FIPS data.
3. Adjust the GeoWorkSpace Coordinate System to match that of the ITD database.
4. Create one new access warehouse, which will serve as the work database.
5. Make sure the Coordinate System of the work database is the same as that of the ITD database.
6. Import all the tables in the conflat_support_tables database into the work database.
7. Import the cityLimit table. Make sure the name of that table in the work database is "cityLimit".
8. Import the urban boundary table. Define one spatial filter according to range of the ITD data (usually, the urban boundary table). Set it as the default spatial filter for any other connection.

9. Import the "Highways" table or all the required tables in the ITD MGE database into the work database using the urban boundary filter. Show them on the map window.

10. Import the TIGER table that contains the required data into the work database (The spatial filter will limit the data imported into the work database). Show it on the map window.

11. Save the workspace.

12. Install the GeoMedia Spatial tools into the GeoMedia professional application

The tool can now be used.

Working with the tool:

1. Launch this tool from the menu. The window will show all of the current connections and lets you choose the connection that connects to the work database.

2. When you click on one connection, another window will show up. There is one combo box in this window. You need to choose the target table you want to work on. There are two target tables. One is B16RoadwayNames, which is final table we will create, the other is RoadwayFromITD, which is the combined table that will be created from the ITD MGE database, if the "Highways" table is not available.

3. If the "Highways" table is not available, choose the RoadwayFromITD. The ITD tables' combo box will then show up. You will need to choose one table from it and click the make button. You choose them one by one until nothing remains in list of combo box.

4. When you finish creating the RoadwayFromITD or you have the "Highways" table, you choose the B16RoadwayNames from the target table combo box. The

window that will show up will display the precise record set combo box, matched record set combo box, distance text box, city name combo box, recalculate all check box, the draw buffer zone command button, the match button, and the make button.

5. There are three functions that can be used. First is the draw buffer zone, which gives the visual effect. Second is the "match", which matches the TIGER data with the ITD data. Third is "make", which makes the B16 table according to the information that "match" creates.

6. If you want to implement the draw buffer zone, you need to choose the precise record set, which is the Higways table or RoadwayFromITD table, and the distance. Click on the draw buffer zone, the tool will draw the buffer zone for you. You can see the relationship between the TIGER data and ITD data visually. Before you save the workspace, you need delete the legend for the buffer zone. The buffer zone is automatically saved in the workdatabase. It can be displayed anytime.

7. If you want to implement the match function, you need to choose the precise record set, match record set, and distance. The precise record set is the Highways table (or the RoadwayFromITD table), and the match record set is the TIGER table.

   To match all the records as soon as possible, you need to run this function many times. The distance needs to be chosen very carefully. Initially you should choose a small value for the distance, say 1 meter, and increase it gradually. You can also choose a range of distance, say 1-10; the tool will run 10 iterations, increasing the buffer distance every time by one meter.

   After each iteration, the matched TIGER records will be deleted from the TIGER table, so you can see the number of TIGER records that still remain. The match information is stored in the precise table, and you can view and modify it as needed.

For modification, you just need to modify the matchedID field. If the TIGER road segment remains on the map but you have already added its ID into matched_Ids field, you should delete it from the matched table. These Ids that stored in matched_Ids field are the primary key of matched table, which is the value of GAVPrimaryKey field in matched table.

The precise table has a distance field that records the distance at which an ITD record got its first matched TIGER record. This information helps you to figure out any match mistakes. There are also conflict_P and conflict_S fields that indicate if the matched TIGER records have conflicts on prefix or suffix. In addition, the result after every iteration is stored in the log table, you can view them. In the log table there is one remark field. You can put your comment related to any problems you found and information for manual modification in this remark field.

8. When all ITD records have matched with the TIGER records, you run make to create the B16, B01, B05, and B11 tables. Before you run make, you need to import a new TIGER table that will be the matched record set and make sure there is one city limit table whose name is cityLimit.

   Usually, you run the make function just once at the end. The tool will use the information that is created by the match function and create the B16 table. If you make some mistake, you can check the recalculate all check box and click the make button, the tool will make the table again. (But beware that if you do this, you need to correct the NewSegmentStart field in the city_FIPS table of the work database. The value of this field should be equal to that of the city_FIPS table in "conflat_support_tables" database).

   If the Done field is true for a record, make doesn't create a new record again. So, if just a few records need to be updated, you can change the value of the Done field in B16 table to false. (One point to note here is that there is no B16RoadwayNames table in the work database. Instead of it, there are two tables, B16RoadwayNames_Highways and B16RoadwayNames_RoadwayFromITD tables. If the precise record set is Highways, the tool will fill the B16RoadwayNames_Highways table, otherwise, the tool will fill the B16RoadwayNames_RoadwayFromITD.)

After you run the make function once, the complete TIGER table will be modified. The new field Highways and RoadwayFromITD are added. These new fields store the ID of the records to which the TIGER record was matched. That helps you to figure out which record in Highways matched with which record in TIGER.

Note that the records in the city_FIPS table in the work database always indicate which segment codes have not been used after the B16 table is created.

After the process is done, you need to do some manual work.

A. Fill the SeqNo field in B16 table manually.

B. Modify the JurisdictionLt and JurisdictionRt fields in B11 table. (Just those records, which is on the city boundary).

9. After the B16 table is created, you are ready to create the intersection table, which is required for creating the C01 MP table. You need to use the GeoMedia Spatial Intersection tool to do this. When using the Spatial Intersection tool, both of the tables we need to choose are B16 table. Make sure the "Output features as" option is Feature class (static, editable), the connection is work database, and the Feature class is intersection. (The feature class name is actually the table name, which contains the result).

10. After you have created the intersection table in the work database, you are ready to click on the makeC01 button.

Advice:

Before running the match command, it is recommended that the TIGER records that do not obviously match to any ITD record be deleted. This will speed up the processing and also help to prevent errors.

When the match function is run, it is better to fix the wrong match as soon as possible. The key iterations are from 1 m to 10 m. Run match for 1 m first, check and correct the mistake, and increase the buffer distance by 1 m for the next iteration, and so on. After every iteration, check for mistakes and correct them. After 10 m, run match for ranges from 11 – 20, 21 – 30, 31 – 40, 41 –100, and 101-150 m. After 150 m, manual matching is advised.

REFERENCES

1. *Milepost And Coded Segment / Roadway Segment Database System.* Administrative Policy A-09-12, Idaho Transportation Department, September 1993.

2. *The MACS/ROSE Access Guide: How to Use the MACS/ROSE Linear Referencing Syste*m. GIS Section, Planning Division, Idaho Transportation Department, November 1997.

3. *Redistricting Census 2000 TIGER/Line ® Files*. U.S. Census Bureau, Washington DC, 2000

4. *NENA Recommended Formats & Protocols For Data Exchange*. National Emergency Number Association, Coshcoton, Ohio, May 1999.

5. *Working with GeoMedia Professional DJA 080740.* Intergraph Corporation, Huntsville, Alabama, 2000.